

## SOCIAL MEDIA ANALYSED

*RepKnight was built to assist Government and commercial intelligence analysts to identify threats and emerging events whilst providing immediate visibility on key metrics that ensure that decision makers have a clear picture of what is happening on the social web. The sheer volumes of messages and content shared during high risk events and incidents would normally make the information unmanageable in any crisis situation. We provide analysts with a quick and easy way to identify the important messages and content and discard the noise, thus making sense of the huge volumes of data and providing a valuable additional source of intelligence to inform operational decisions.*

**John Reid - CEO RepKnight**



## RepKnight: Delivering Intelligence in Real Time

### A User Focused Application

Simple, yet a beautiful and intuitive user experience

At RepKnight, user experience is paramount. We believe that an enhanced user experience drives data discovery, putting the identification of the “Needle in the Haystack” within the reach of experienced open source intelligence analysts. Highlighting of critical information / actionable insights on a regional and global basis. Beautiful data-visualisation assists intelligence analysts to get a view at a glance, whilst also acting as a translation mechanism to management.

### Data On A RepKnight Scale

We specialise in firehose and large volume data

RepKnight is the fastest open source intelligence platform within the global marketplace. We provide both sentiment and geographical analysis in real time, providing an alert when there is a spike in the data volumes. Our system is built to specifically deal with the high volume search terms, and sharp spikes in the data volumes.

### Make Sense Of Your Data

Real-time charting that helps you analyse results

RepKnight's unique platform ensures that analysts get visibility of the data relevant to them faster and in a more user friendly format. Our tool provides enhanced user experience whilst still providing senior management with the hard numbers, graphs and piecharts. RepKnight indexes data from Twitter, Facebook, Google+, YouTube, Flickr, News Sites and Forums. You can also add in source sites of your own, allowing you to build a bespoke search engine.

To arrange a presentation of the RepKnight Software please contact John Zeledon at [jzeledon@ADITechnologies.com](mailto:jzeledon@ADITechnologies.com)

# Predicting Cyber Attacks through Interaction and Actor Behavior Modeling and Event Detection in Dark Web, Black Market and Underground Forums

Battelle Cyber Innovations  
solutions@battelle.org

## ABSTRACT

Current state-of-the-art cyber security technology relies heavily on signatures to derive threats or anomalies. While this approach has proven valuable in the past, attacks have grown in sophistication and these techniques have led to cyber security practitioners to focus on the effects of an attack as opposed to determining and mitigating the cause. A shortcoming of these systems is a lack of novel sources for data enrichment and probabilistic warnings based on unconventional data sources. One such source is the significant amount of cyber threat activity that occurs within Dark Web, black market and underground forums/marketplaces. However, unlike traditional social media, forum data presents problems when performing social network analysis and event detection. Properly modeling interactions among nefarious actors can lead to accurate depictions of threat community behavior. This technique aids in uncovering illicit networks, identifying influential members, and generating accurate social graphs.

In its proposed research for CAUSE, Battelle will seek to enhance its technology called DarkScout that collects and integrates forum, marketplace and social content with flexible tools for the structured consumption of irregular media in hostile web environments. DarkScout uses language-agnostic algorithms to collect, organize and analyze contextual information surrounding media items to uncover community structures, and adversary pattern-of-life, trends, motivation, intent and capabilities. Further, SME-developed ontologies provide a foundation for performing event detection and training text analysis classifiers to extract potential indicators of cyber attack. The identification of anomalous events is augmented with pattern-of-life analysis to provide a temporal view of incoming threats.

In its CAUSE-supported research, Battelle will augment Interaction Modeling, Actor Behavior Models and integrate technology to analyze and de-anonymize Bitcoin sale/purchase activity to capture communication exchanges more accurately within threat-actor forums and enrich it with temporal event data, yielding robust information propagation models. Propagation models provide definitive analysis of how far and how fast information spreads, and indicates threat momentum. Entity extraction of technical indicators will inform propagation models to provide a realistic view of whether threat actors are seeking to exploit an attack vector, and provide early warning of cyber attacks via a robust Application Programming Interface. Battelle will seek to work with other researchers who are developing advanced Intrusion Detection Systems and network sensor systems that would benefit from enriched data sources and models depicting threat-actor behavior and activity.

**Dr. Paulo Shakarian** is an Assistant Professor at Arizona State University where he works in the Big Data group. He specializes in advanced data analytics, network science, artificial intelligence, and cyber-security. Specific application domains have included intelligence analysis, counter-insurgency, counter-IED, law-enforcement, and cyber-security. His previous work has been presented at major academic venues including KDD, AAMAS, and ESORICS as well as industry conferences such as ShmooCon. His work has been funded by the ARO, DARPA, IARPA, and USAF A2II. Shakarian's work on analyzing geospatial data resulted in the "SCARE" software for locating weapons caches that was used by Task Force Paladin in Afghanistan and also featured in *The Economist*. His work on social network data analytics resulted in the "GANG" and "SNAKE" software packages that are currently in use by the Chicago Police and also featured in *Popular Science*. Dr. Shakarian is also the author of two books, including Elsevier's *Introduction to Cyber-Warfare*. Previously, Dr. Shakarian was a commissioned officer in the U.S. Army where he worked in a variety of intelligence positions that include combat tours in Operation Iraqi Freedom. He is a recipient of the Bronze Star and Army Commendation Medal for Valor.

Contact information:

## **Paulo Shakarian, Ph.D.**

**Assistant Professor and**

**Director, Cyber-Socio Intelligent Systems (CySIS) Lab**

**Arizona State University**

BYENG 408

699 S. Mill Ave.

Tempe, AZ 85281

(480) 727-5290 (o)

(480) 965-2751 (f)

[shak@asu.edu](mailto:shak@asu.edu)

<http://shakarian.net/paulo>

**Florida Center for Cybersecurity (FC<sup>2</sup>)  
University of South Florida (USF)  
Adaptive Immersion Technologies (AIT)**



The Florida Center for Cybersecurity (FC<sup>2</sup>), with ongoing funding from the Florida State Legislature, is an initiative to lead and coordinate cybersecurity research, education and outreach across the state and beyond. FC<sup>2</sup> brings access to researchers across all 12 institutions in Florida's State University System (SUS) as well as collaborative engagement among government, defense and business communities.

The University of South Florida (USF) is a high-impact, global research university, classified by the Carnegie Foundation for the Advancement of Teaching as a community engaged university and as a the top tier of research university, a distinction attained by only 2.2 percent of all universities. USF is home to the Program in National and Competitive Intelligence, an Intelligence Community Center of Academic Excellence. USF and FC<sup>2</sup> also have been designated by the National Security Agency (NSA) and the Department of Homeland Security (DHS) as a National Center of Academic Excellence in Information Assurance/ Cybersecurity for academic years 2014-2019. USF's School of Information has distinctive capabilities and subject matter expertise in (1) Cyber Intelligence – collection and analysis of information concerning the intentions, capabilities, activities of adversaries and competitors in the cyber domain; (2) Applying intelligence analytic methods and technologies to the cyber domain; (3) Adversary characterization, risk/threat assessment, and threat actor behavior; and (4) Integrating collection and knowledge management technologies to improve the efficiency of cybersecurity operations. Key personnel include “Scuba” Steve Gary – Former Chief of Cyber Intelligence, US Special Operations Command, who holds an active clearance and an MS in Cyber Operations and; Dr. Randy Borum – (Cleared) Psychologist and former science advisor to the DNI and IARPA who specializes in applying intelligence analytic methods to the cyber domain, strategy and analytic decision making.

Adaptive Immersion Technologies (AIT) develops simulation-based personnel selection, training, and performance management systems that promote human resilience in physically, cognitively, and psychologically taxing performance domains. They tailor systems to integrate a trilogy of performance solutions through prediction, enhancement, and support, using a synthesis of concepts from data science, human simulation, and adaptive assessment technology to optimize human performance. AIT has distinctive capabilities and subject matter expertise in (1) Novel machine learning applications to complex human performance prediction problems; (2) Computational modeling of human performance; (3) Technology-enabled training, performance assessment and diagnosis employing modern psychometric theory; and (4) Algorithm development, optimization, and benchmarking for real time, simulation-based assessment.

Contact:  
“Scuba” Steve Gary  
Assistant Professor of Practice  
School of Information, University of South Florida  
[sgary@usf.edu](mailto:sgary@usf.edu)

## **Galois and Adversarial Reasoning**

Galois, Inc. is interested in applying our experience in the area of adversarial reasoning to the IARPA CAUSE program. The Adversarial Reasoning effort at Galois comprises three complementary threads:

1. Investigating the nature of deception and counterdeception, particularly as it applies to the cyber domain. Cyber adversaries rely on deceptive attack techniques, and understanding patterns of deception enables accurate predictions and proactive counterdeceptive responses.
2. Developing strategic, cognitive and game-theoretic approaches to developing cyber actor models, with a focus on understanding how to reason about the beliefs, intentions, and objectives of cyber adversaries.
3. Formulating techniques that allow for robust reasoning under conditions of extreme uncertainty and ambiguity, especially in those circumstances where statistical data is absent, and the only evidence available is likely to be fragmentary or conflicting.

We have both prior and ongoing unclassified work with DOD agencies to research and prototype these capabilities. For the IARPA CAUSE opportunity, we are interested in partnering with other performers with expertise and interest in any of the following areas:

- Methods to manage and extract huge amounts of streaming and batch data
- Development of models to generate probabilistic warnings of future cyber events
- Multiple sensors not typically used in the cyber domain

## **About Galois**

Galois, Inc. is a computer science R&D firm with the mission to “create trustworthiness in critical systems.” Founded in 1999, and located in Portland, Oregon, Galois applies cutting-edge computer science and mathematics to solve difficult technological problems.

Over its 15-year life, Galois has worked to bring rigorous, mathematically based techniques to challenges in domains such as software correctness, cryptography, cyber-physical systems, mobile security, machine learning, and human computer interaction. Galois continues to work with a wide variety of government and commercial clients, particularly in the DOD and IC.

IBM's Cognitive Cyber Security Defense (CCD) is a big data and analytic solution that employs machine learning techniques to provide an adaptive and agile defensive posture in real-time. It is an integrated solution with proven machine learning models from IBM Research with the ability to build new families of Cyber models to react to the ever changing Advanced Persistent Threat (APT) environment. The Cognitive Cyber Security Defense solution is designed to scale from an entry-basic configuration up to a full-capability system depending on your cyber defense needs. It provides a machine learning workbench for the development of your own predictive cyber models. These models can perform behavior analytics as well as target specific DNS related attack types as well as behavior modeling of netflow data. Key attributes of the system are:

1. The CCD solution is an APT detector comprised of a family of pretrained machine learning models outputted to rich visualization
2. Solution runs on x-86 infrastructure running RHEL 6.1 or higher
3. It can connect to existing Cyber SIEM, Big Data or Cloud Solutions
4. The Models dynamically update with changing threat vectors
5. We have field tested this solution with numerous customers from Utilities to telcos to a large commercial entities

## INTELLIGENT SYSTEMS FOR FORECASTING AND DETECTING CYBER ATTACKS

OUR EXPERTISE AND PROVEN CAPABILITIES	DECISION SUPPORT	BEHAVIOR MODELING	CYBER SIMULATION
	<b>Multi-Future Probabilistic Forecasting</b>	<b>Cognitive Red-Team Agents</b>	<b>Cyber Sandbox for Attack &amp; Defense Training</b>
<b>AUTONOMY</b>	Multiple agents search alternative paths through complex behavior models in parallel	Enables scalable, repeatable wargaming and testing of security and network infrastructure	Agent-based cyber ecology provides autonomous adversaries and legitimate users
<b>ADAPTATION</b>	Any-time forecast adjusts in real time to incoming data; runs 10 <sup>4</sup> x faster than real time	Learned behavior model exposes novel attack vectors	Dynamic Tailoring adapts adversaries and environment to maximize learning
<b>HUMAN/SYSTEM INTERFACE</b>	Yields probability distribution over alternative futures to support ACH and mitigate cognitive anchoring	Agents acting as virtual assistants allow human experts to focus on high level goals	Constructive sims are readily available for continuous training and evaluation

Figure 1: Matrix of SoarTech's core capabilities, related applications and relevant examples.

**The Challenge:** Current approaches to detecting cyber attacks are *reactive* and *shallow*. They are *reactive* because they focus on what adversaries have done in the past, rather than anticipating what they may do in the future. They are *shallow* because they focus on cyber observables without reasoning about adversaries' goals and objectives.

**Potential Approach:** Applying SoarTech's core capabilities (shown in Figure 1) we have experience to combine innovative analytics on observables with sophisticated behavioral models of cyber actors that can support both cognitive reasoning (extending the model through experience and explaining reasoning to humans) and Monte Carlo exploration (for probabilistic forecasting over multiple possible futures).

## ENHANCING CYBERSPACE DEFENSE THROUGH BIDIRECTIONAL BEHAVIORAL MODELS

### SoarTech POCs:

Dylan Schmorrow, Ph.D.  
Chief Scientist  
703.424.3138

[dylan.schmorrow@soartech.com](mailto:dylan.schmorrow@soartech.com)

Denise Nicholson, Ph.D., CMSP  
Director of X  
407.616.7651

[denise.nicholson@soartech.com](mailto:denise.nicholson@soartech.com)

H. Van Dyke Parunak, Ph.D.  
Senior Scientist  
734.395.3253

[van.parunak@soartech.com](mailto:van.parunak@soartech.com)

We are interested in discussing partnerships and collaborations.

Below are the logos of a few of our Sponsors and Partners for related research.





## **Cyber-attack Automated Unconventional Sensor Environment (CAUSE)**

**SRA International, Inc.**

**Joseph Pemberton**

**[joe\\_pemberton@sra.com](mailto:joe_pemberton@sra.com)**

**703-803-1882**

### **Abstract**

The growth of cyber incidents, from identify theft to full-scale attacks on corporate and Government cyber assets, highlights the importance of cyber defense to the economic and political security of the country. Existing cyber defense methods typically focus on forensic assessments to answer the question, “what happened?” rather than attempting to predict cyber-attacks before they occur. SRA is actively researching ways to improve cyber defense capabilities for our customers. We have extensive experience providing cyber security services to a wide variety of Government agencies. Our subject matter experts understand the current cyber defense landscape, and we have experience and relationships with a wide range of cyber security solution vendors.

SRA has developed and markets NetOwl®, a suite of text and entity analytics products. We have recently expanded NetOwl’s ontology to cover counterterrorism, intelligence, military, homeland security, law enforcement, business, compliance, and cyber security related areas. Our cyber security ontology integrates cyber-event concepts from the U.S. Department of Defense, US-CERT, and other cyber security organizations, and terminology for critical infrastructure such as energy, financial, and telecommunications facilities and organizations – prime targets of cyber-attacks that could affect U.S. security. Our expanded cyber security ontology is designed to allow organizations to process large volumes of unstructured content and automatically identify key cyber-related events as well as entities involved in these cyber events

Our team includes Context Relevant which provides an optimized machine learning pipeline for analyzing complex data sets. Context Relevant software helps data scientists process unstructured (e.g., dirty/messy) data, including Open Source data, process millions of input dimensions, and automatically general and explore hundreds of thousands of models in parallel. Their tools have been used to general efficient prediction of future events. With our teammates, we are currently investigating the combination of NetOwl, deep machine learning methods, and big data processing to the problem of analysis and prediction of cyber-attacks. In particular, we are investigating ways to combine both structured and unstructured (messy) data from traditional and nontraditional sources to enable early stage detection of cyber-attacks.

SRA is eager to join the IARPA CAUSE Program, and we are looking for additional teammates to help round out our team.



## About ViON

- ViON is a veteran-owned, privately held company with over 34 years of experience building IT enterprise solutions for government and commercial customers. Being independent allows for streamlined decision making and nimble responses to our customers' needs.
- ViON works with the largest and most innovative OEM suppliers in the industry to design and implement solutions that meet any IT storage or server need. Partners include IBM, Brocade, Cisco, EMC, Hitachi Data Systems, NetApp, and many more.
- Known for our engineering expertise and exacting standards, our team ensures that only those with the highest level of training, experience, and industry certifications design, install, maintain, and support our breadth of solutions.

## Identity Analytics for National Security

Our modern world is a place where national security, critical infrastructure and national resources may be at risk due to the actions of small groups of people with anti-social agendas. Those with malicious intent are typically highly motivated to operate covertly, going to great lengths to obscure their identities, relationships and organizational affiliations.

A key weapon in the fight against crime, civil unrest and terrorism is timely and actionable information. Better recognition of identities, relationships and affiliations can provide the hidden intelligence needed to anticipate and prevent catastrophe. To harness data from all available sources and extract actionable intelligence demands unique software-driven capability. ViON's Data Adapt Analytics A2000 appliance for Threat Prediction and Prevention, based upon the Cisco UCS platform, is powered by IBM software specifically designed to recognize attempts to obscure identities and relationships.

## The Data Adapt Analytics Solution

The ViON Data Adapt Analytics A2000 appliance utilizes proven entity resolution and analysis capabilities that have been used for years within government organizations that protect national security. These mission requirements range from borders and port entry to complex counter-insurgency and inside threat detection.

The A2000 was architected to ingest all available data sources. The analytical engine detects intersections in the data that reveal clues about identities and relationships, continuously learning with each new set of data. In this way the data itself reveals the clues that it hides, providing the critical Non Obvious Relationship Awareness. This approach also reveals multiple degrees of separation between people, essential for mapping out social networks. Behavioral and pattern-based analytics create the ability to observe coordinated movements that would otherwise avoid detection. These analytics are configurable based on thresholds and risk scores that determine who is alerted, when and how.

By integrating all these capabilities into a scalable x86 hardware platform, ViON has created a new generation of solution that delivers the value of a complex custom enterprise software system but with the same speed, ease and economy of an appliance.

## Mission Optimized Threat Prediction & Prevention Capabilities

The system's sophisticated disambiguation technologies were specifically created to penetrate the cultural ambiguities, fabrications & identity misrepresentation tactics nefarious groups and individuals use to hide. It uncovers and links obvious and non-obvious relationships which may reveal criminal syndicates, gangs, revolutionary organizations or terrorist cells. Capabilities include:

- **Full attribution:** Achieve greater accuracy with an entity resolution and analytic engine that accumulates a history as opposed to a snapshot of each individual or company in the database over time.
- **Relationship resolution:** Identify & compare non-obvious relationships between addresses, phone numbers, e-mail addresses, and other characteristics discovered and linked across multiple individuals
- **Link analysis:** Analyze, visualize, and extend these relationships building complex structures, revealing the hierarchies and methods of operation employed by criminal, terrorist and fraudulent networks
- **Real-time changes:** Compare identity records to the database upon receipt to determine if it resolves an existing record, is new, or requires and unresolve of an existing record.
- **Self-healing and self-correcting:** Automatically examine and update any entity in the repository that would be affected based on new observations.
- **Autonomic real-time alerting:** Automatically check against existing information and generate alerts allowing for detailed analysis of involved parties as new data is ingested.
- **Global name resolution:** Apply linguistic rules automatically to find matches through cultural context via patented IBM technology.
- **Behavioral and pattern analysis:** Uncover coordinated activities and patterns that provide the capacity to anticipate a pending threat event via deep data mining and statistical analysis.

## Why Data Adapt Analytics Solutions?

More than simply bundling hardware and software, the secret to Data Adapt Analytics offerings such as Threat Prediction & Prevention is purposeful design. Lessons learned from numerous prior projects have been incorporated into the product so that customers can begin with configuring the system to their environment, data and mission rather than locating and hiring the highly specialized skills needed to operationalize the different software components. Integrated technologies and accelerators are combined in different ways unique to the characteristics of a given mission. The application exploits the right mix of system resources (processing power, memory and storage, for example) for deeper levels of optimization to achieve the desired scale, performance, analysis, and user experience.

To learn more of how the ViON Data Adapt Analytics solutions can help you, please visit [www.vion.com](http://www.vion.com).



196 Van Buren Street | Herndon, Virginia 20170  
(571) 353-6000 | (800) 761-9691 | [vion.com](http://vion.com)

